

# Generating Product Systems

Nir Avni\*, Benjamin Weiss†

July 7, 2009

## Abstract

Generalizing Krieger's finite generation theorem, we give conditions for an ergodic system to be generated by a pair of partitions, each required to be measurable with respect to a given sub-algebra, and also required to have a fixed size.

## 1 Introduction

Let  $(Z, \mathcal{F}_Z, \mu)$  be a probability space, and let  $T : Z \rightarrow Z$  be a measurable, probability preserving transformation of  $Z$ . If  $\mu$  is ergodic with respect to  $T$ , we say that  $(T, \mathcal{F}_Z, \mu, T)$  is an ergodic system.

A partition  $P$  of  $Z$  is called generating if the sigma-algebra generated by  $\bigvee_{-\infty}^{\infty} T^i P$  and the sigma-algebra of null and co-null sets is equal to  $\mathcal{F}_Z$ . If  $P$  has  $a$  parts, then for any  $\bigvee T^i P$ -measurable partition  $Q$ , the entropy  $h(T, Q)$  is less than or equal to  $\log a$ . Hence, if  $h(T) > \log a$ , then there are no generating partitions with  $a$  parts. In [4], Krieger proved the following partial converse:

**Theorem 1.1** (Krieger). *Let  $(Z, \mathcal{F}_Z, \mu, T)$  be an ergodic system, and let  $a$  be an integer such that  $\log a > h(T)$ . Then there is a generating partition  $P$  of  $Z$  that has  $a$  parts.*

The case where  $h(T)$  is equal to  $\log a$  is different; there is a unique ergodic system for which there is a generating partition with  $a$  parts (namely, the Bernoulli system).

Krieger's theorem can be thought of as an infinite version of Shannon's coding theorem for noiseless channels. In Shannon's theorem the block size is finite and the error remains positive but converges to zero as the size of the block grows. Shannon's theorem for more general channels was also given a zero error version using infinite codes, as is usual in ergodic theory in a series of papers including [2] and [3].

In information theory there is a surprising extension of Shannon's coding theorem to correlated sources due to Slepian and Wolf, [6]. Our main

---

\*Department of Mathematics, Harvard University, Cambridge MA 02138. Email address: avni.nir@gmail.com

†Institute of Mathematics, The Hebrew University of Jerusalem, Jerusalem 91904 Israel. Email address: weiss@math.huji.ac.il

result is a zero error version of the Slepian–Wolf result for noiseless channels. Combining the techniques of Kieffer with those of this paper would lead to a zero error version of the Slepian–Wolf theorem in the context of noisy channels, but we leave these extensions for the future.

In order to describe the result, consider the following scenario: Suppose that  $\mathcal{F}_X, \mathcal{F}_Y \subset \mathcal{F}_Z$  are sub-sigma-algebras and that  $\mathcal{F}_Z = \mathcal{F}_X \vee \mathcal{F}_Y$  (without loss of generality, we may assume that  $Z = X \times Y$  is a product space, and that  $\mathcal{F}_X, \mathcal{F}_Y$  are the corresponding sigma-algebras). We seek a pair of partitions  $P_X, P_Y$  such that  $P_X$  is  $\mathcal{F}_X$ -measurable and has  $a$  parts,  $P_Y$  is  $\mathcal{F}_Y$ -measurable and has  $b$  parts, and the partition  $P_X \vee P_Y$  generates. There are three necessary conditions for such partitions to exist: the first is that the conditional entropy of  $T$  with respect to  $\mathcal{F}_Y$ , which we denote by  $h(T|\mathcal{F}_Y)$ , is less than or equal to  $\log a$ , the second condition is that  $h(T|\mathcal{F}_X) \leq \log b$ , and the third condition is that  $h(T) \leq \log a + \log b$ . As in Theorem 1.1, we show that these conditions are almost sufficient:

**Theorem 1.2.** *Let  $(Z, \mathcal{F}_Z, m, T)$  be an ergodic and probability preserving system, let  $\mathcal{F}_X, \mathcal{F}_Y$  be invariant sub-sigma-algebras of  $\mathcal{F}_Z$ , and let  $a$  and  $b$  be integers. Assume that  $\log a > h(T|\mathcal{F}_Y)$ , that  $\log b > h(T|\mathcal{F}_X)$ , and that  $\log a + \log b > h(T)$ . Then there are partitions  $P_X$  and  $P_Y$  of  $Z$  such that*

1.  $P_X$  is measurable with respect to  $\mathcal{F}_X$  and has  $a$  parts.
2.  $P_Y$  is measurable with respect to  $\mathcal{F}_Y$  and has  $b$  parts.
3.  $P_X \vee P_Y$  generates the sigma-algebra  $\mathcal{F}_X \vee \mathcal{F}_Y$ .

More generally, it is straightforward to generalize Theorem 1.2 to the case of  $n$  partitions. Namely, assume that there are  $n$  sigma-algebras,  $\mathcal{F}_1, \dots, \mathcal{F}_n$ , and  $n$  integers  $a_1, \dots, a_n$ . If for every non-empty subset  $S \subset \{1, \dots, n\}$ ,

$$h\left(T \mid \bigvee_{i \notin S} \mathcal{F}_i\right) < \sum_{i \in S} \log a_i,$$

then there are partitions  $P_1, \dots, P_n$  such that  $P_i$  has  $a_i$  parts and is  $\mathcal{F}_i$ -measurable, and that  $P_1 \vee \dots \vee P_n$  generates.

The proof of Theorem 1.2 is given in Section 4, and consists of two parts. In the first part (Proposition 4.2) we show that there are partitions  $P_X, P_Y$  that satisfy the first two requirements, and such that the partition  $P_X \vee P_Y$  is close to being a generating partition (in the sense that the sigma-algebra generated by it contains a partition which is close to a fixed generating partition of  $Z$ ). This is done by a procedure of ‘painting names along Rohlin towers’. In the second part (Proposition 4.4) we show that one can slightly change these partitions and get partitions that are even closer to being generating. The proof of Proposition 4.4 uses the procedure of ‘re-painting names along Rohlin towers’. Applying Proposition 4.4 inductively, we get a Cauchy sequence of partitions that converges to a generating partition, which proves Theorem 1.2.

The procedures of painting and repainting along Rohlin towers are explained in Section 3. Finally, Section 2 contains several auxiliary results, the important one is a corollary of the Shannon–McMillan theorem. No novelty is claimed for this material; it is assembled for the reader’s convenience.

## 2 Preliminaries

**Definition 2.1.** Let  $(Z, \mathcal{F}_Z, \mu, T)$  be an ergodic system.

1. A (Rohlin-) tower is a pair  $\mathcal{T} = (A, M)$ , where  $A$  is a subset of  $Z$ , and  $M$  is a natural number, such that for any  $0 < i < M$ , the sets  $A$  and  $T^i A$  are disjoint.
2. If  $\mathcal{T} = (A, M)$  is a tower, the set  $A$  is called the base of the tower, and is denoted by  $b\mathcal{T}$ ; the number  $M$  is called the height of the tower, and is denoted by  $|\mathcal{T}|$ .
3. If  $\epsilon$  is a real number, we say that a tower  $\mathcal{T} = (A, M)$  covers more than  $\epsilon$  of  $Z$  if  $\mu(A \cup TA \cup \dots \cup T^{M-1}A) > \epsilon$ .
4. If  $\mathcal{G}$  is an invariant sub-sigma-algebra, we say that  $\mathcal{T}$  is  $\mathcal{G}$ -measurable if  $A$  is  $\mathcal{G}$ -measurable.

**Definition 2.2.** Let  $(Z, \mathcal{F}_Z, \mu, T)$  be an ergodic system, and let  $k$  be an integer. A partition of  $Z$  into  $k$  parts is a function  $P : Z \rightarrow \{\overline{0}, \dots, \overline{k-1}\}$ . Given a partition  $P$  into  $k$  parts, we denote the number  $k$  by  $|P|$ .

**Definition 2.3.** Let  $(Z, \mathcal{F}_Z, \mu, T)$  be an ergodic system, let  $\mathcal{T} = (A, M)$  be a tower, and let  $P$  be a partition of  $Z$ . For every  $z \in A$ , the  $(\mathcal{T}, P)$ -name of  $z$  is the tuple  $(P(z), \dots, P(T^{M-1}z))$ . We denote the  $(\mathcal{T}, P)$ -name of  $z$  by  $\nu_{\mathcal{T}, P}(z)$ .

For a proof of the following theorem, see [5].

**Theorem 2.4.** Let  $(Z, \mathcal{F}_Z, \mu, T)$  be an ergodic system, let  $f : Z \rightarrow \mathbb{C}$  be a bounded function, and let  $P$  be a partition of  $Z$ . For every  $\epsilon > 0$  there is a natural number  $N$  such that if  $\mathcal{T}$  is a tower of height greater than  $N$  that covers more than  $\epsilon$  of the space, then

1. The set of points  $z$  in the base of  $\mathcal{T}$  that satisfy

$$\left| \frac{1}{|\mathcal{T}|} \left( f(z) + f(Tz) + \dots + f(T^{|\mathcal{T}|-1}z) \right) - \int_Z f d\mu \right| < \epsilon$$

has measure greater than  $(1 - \epsilon)\mu(A)$ .

2. There is a collection  $\mathcal{N}$  of  $2^{(h(Z, P) + \epsilon)|\mathcal{T}|}$  elements in  $|P|^{|\mathcal{T}|}$  such that
  - (a) For every  $n = (n_i) \in \mathcal{N}$ , the measure of the set of  $z$ 's in  $A$  whose  $(\mathcal{T}, P)$ -name is equal to  $n$  is between  $2^{-(h(Z, P) + \epsilon)|\mathcal{T}|}$  and  $2^{-(h(Z, P) - \epsilon)|\mathcal{T}|}$ .
  - (b) The measure of the set of  $z$ 's in  $A$  whose  $(\mathcal{T}, P)$ -name is in  $\mathcal{N}$  is greater than  $(1 - \epsilon)\mu(A)$ .

**Definition 2.5.** For a partition  $P$ , define  $\mathcal{N}_P$  to be the set of words of unspecified length with entries in  $\{\overline{0}, \dots, |P| - 1\}$ .

For a set  $\Omega$ , we denote by  $2^\Omega$  the collection of finite subsets of  $\Omega$ .

**Corollary 2.6.** Let  $(Z, T)$  be an ergodic system, and let  $P, Q$  be two partitions of  $Z$ . Denote the sub-algebra  $\bigvee_{i=-\infty}^{\infty} T^i P$  by  $\mathcal{F}$ . For every  $\epsilon > 0$ , there is an  $N$  such that if  $\mathcal{T} = (A, M)$  is a tower of height greater than  $N$  that covers more than  $\epsilon$  of the space, then there is a function

$$\Phi : \mathcal{N}_P \rightarrow 2^{\mathcal{N}_Q}$$

such that

1. For every  $z \in A$ , the set  $\Phi(\nu_{T,P}(z))$  has less than  $2^{|\mathcal{T}|(h(T,Q|\mathcal{F})+\epsilon)}$  elements.
2. The set of  $z$ 's in  $A$  such that  $\Phi(\nu_{T,P}(z))$  contains the  $(T, Q)$ -name of  $z$  has measure greater than  $(1 - \epsilon)\mu(A)$ .

*Proof.* By Theorem 2.4, if  $N$  is large enough, then there are collections

$$\mathcal{N} \subset |P|^{|\mathcal{T}|} \quad \mathcal{M} \subset |P \times Q|^{|\mathcal{T}|}$$

such that for any  $n \in \mathcal{N}$ , the measure of the set of points whose  $(T, P)$ -name is equal to  $n$  is less than  $2^{-|\mathcal{T}|(h(T,P)-\epsilon)}$ , for every  $m \in \mathcal{M}$ , the measure of the set of points whose  $(T, P \times Q)$ -name is equal to  $m$  is greater than  $2^{-|\mathcal{T}|(h(T,P \times Q)+\epsilon)}$ , and the set of points whose  $(T, P)$ -names are not in  $\mathcal{N}$  or whose  $(T, P \times Q)$ -names are not in  $\mathcal{M}$  has measure less than  $\epsilon\mu(A)$ .

Let  $\pi_1 : |P \times Q|^{|\mathcal{T}|} \rightarrow |P|^{|\mathcal{T}|}$  and  $\pi_2 : |P \times Q|^{|\mathcal{T}|} \rightarrow |Q|^{|\mathcal{T}|}$  be the co-ordinate projections. Let  $n \in \mathcal{N}$ . If  $n \in \mathcal{N}$  then

$$2^{-|\mathcal{T}|(h(T,P)-\epsilon)} \geq \mu\{z \in A | \nu_{T,P \times Q}(z) \in \pi_1^{-1}(n) \cap \mathcal{M}\} \geq |\pi_1^{-1}(n) \cap \mathcal{M}| 2^{-|\mathcal{T}|(h(T,P \times Q)+\epsilon)}$$

and so  $|\pi_1^{-1}(n) \cap \mathcal{M}| \leq 2^{|\mathcal{T}|(h(T,Q|\mathcal{F})+2\epsilon)}$ . Define  $\Phi(n) = \pi_2(\pi_1^{-1}(n) \cap \mathcal{M})$  if  $n \in \mathcal{N}$ , and  $\Phi(n) = \emptyset$  if  $n \notin \mathcal{N}$ .  $\square$

We will use the standard notation for the entropy function:  $H(x) = x \log_2(x) + (1-x) \log_2(1-x)$ . The following is well known. For a (much more) general statement, see [1, Theorem 3.9].

**Lemma 2.7.** *Let  $P$  be a partition, let  $\mathcal{G}$  be an invariant sub-sigma-algebra, and let  $\mathcal{T} = (A, M)$  be a  $\mathcal{G}$ -measurable tower that covers more than  $1 - \epsilon$  of the space. If there is a  $\mathcal{G}$ -measurable function  $\Phi : A \rightarrow 2^{\mathcal{N}_P}$  such that the number of elements in  $\Phi(z)$  is less than  $2^{h|\mathcal{T}|}$  and the set of  $z \in A$  such that  $\nu_{T,P}(z) \in \Phi(z)$  has measure greater than  $(1 - \epsilon)\mu(A)$ , then  $h(T, P|\mathcal{G}) < h + H(\epsilon) + 2\epsilon|P|$ .*

**Definition 2.8.** *Let  $P, Q$  be partitions of a probability space  $(Z, \mu)$ . Define*

$$|P \Delta Q| = \mu\{z \in Z | P(z) \neq Q(z)\}.$$

**Definition 2.9.** *Let  $P$  be a partition, let  $\mathcal{A}$  be an algebra of sets, and let  $\epsilon > 0$ . We write  $P \stackrel{\epsilon}{\subset} \mathcal{A}$  if there is an  $\mathcal{A}$ -measurable partition  $Q$  such that  $|P \Delta Q| < \epsilon$ . If  $Q$  is a partition, we write  $P \stackrel{\epsilon}{\subset} Q$  if  $P \stackrel{\epsilon}{\subset} \mathcal{A}_Q$ , where  $\mathcal{A}_Q$  is the algebra generated by  $Q$ .*

The following lemma is evident:

**Lemma 2.10.** *Let  $P_1, P_2, P_3$  be partitions of a probability space  $Z$ . If  $P_1 \stackrel{\epsilon}{\subset} P_2$  and  $P_2 \stackrel{\delta}{\subset} P_3$ , then  $P_1 \stackrel{\epsilon+\delta}{\subset} P_3$ .*

**Proposition 2.11.** *Let  $(Z, T)$  be an ergodic system, and let  $P, Q$  be partitions of  $Z$ . Assume that  $P \stackrel{\delta}{\subset} \bigvee_{-\infty}^{\infty} T^i Q$ . Then for every  $\epsilon > 0$ , there is an  $K$ , such that for every tower  $\mathcal{T} = (A, M)$ , which covers more than  $\epsilon$  of the space and for which  $M > K$ , there is a function  $\Phi : \mathcal{N}_Q \rightarrow 2^{\mathcal{N}_P}$  such that*

1. For every  $z \in A$ ,  $|\Phi(\nu_{T,Q}(z))| < 2^{(H(\delta)+\delta|P|+\epsilon)M}$ .
2. The set of  $z \in A$  such that  $\nu_{T,P}(z) \in \Phi(\nu_{T,Q}(z))$  has measure greater than  $(1-\epsilon)\mu(A)$ .

*Proof.* By assumption, there is a number  $N$ , and a function  $\phi : |Q|^{2N+1} \rightarrow |P|$  such that the set

$$B = \{z \in Z \mid \phi(Q(T^{-N}z), \dots, Q(T^Nz)) = P(z)\}$$

has measure greater than  $1-\delta$ . Let  $\eta > 0$ , to be chosen later. By Theorem 2.4, if  $T = (A, M)$  is a tower that covers more than  $\epsilon$  of the space, and  $M$  is sufficiently large, then the set of  $z \in A$  such that

$$\frac{|\{0 \leq i < M \mid T^i z \in B\}|}{M} > 1 - \delta - \eta/2$$

has measure greater than  $(1-\epsilon)\mu(A)$ . If we require in addition that  $N < \eta M/2$  we get that for those  $z$ 's,

$$\frac{|\{N \leq i \leq M - N \mid T^i z \in B\}|}{M} > 1 - \delta - \eta. \quad (1)$$

Define  $\psi : |Q|^M \rightarrow |P|^M$  by

$$\psi((q_i))_j = \begin{cases} \phi(q_{j-N}, \dots, q_{j+N}) & \text{if } N \leq j \leq M - N \\ \bar{0} & \text{else} \end{cases}.$$

For  $(q_i) \in |Q|^M$ , let  $\Phi((q_i))$  be the set of elements in  $|P|^M$  whose Hamming distance from  $\psi((q_i))$  is less than  $(\eta + \delta)M$ . The size of  $\Phi((q_i))$  is less than

$$\binom{M}{(\eta + \delta)M} |P|^{(\eta + \delta)M} < 2^{(H(\eta + \delta) + (\eta + \delta)|P|)M},$$

which is less than  $2^{(H(\delta) + \epsilon + \delta|P|)M}$  if  $\eta$  is taken small enough. Moreover, for every  $z$  that satisfies (1),  $\nu_{T,P}(z) \in \Phi(\nu_{T,Q}(z))$ .  $\square$

We end this section by the following lemma, which should be thought of a relative version of the claim that if  $f$  is a random function from a set of size  $a$  to a set of size  $b$  then the non-empty fibers of  $f$  have size  $1 + O(a/b)$ .

**Lemma 2.12.** *Let  $(Z, \mathcal{F}_Z, \mu)$  be a probability space, and let  $\Omega$  be a finite set. Assume that  $\phi : Z \rightarrow \Omega$ ,  $\Phi : Z \rightarrow 2^\Omega$  are functions such that  $|\Phi(z)| < a$  and  $\phi(z) \in \Phi(z)$  for every  $z \in Z$ . The probability that a random function  $\psi : \Omega \rightarrow [b]$  satisfies that*

$$\mu \left\{ z \in Z \mid |\psi^{-1}(\psi(\phi(z))) \cap \Phi(z)| < 1 + \frac{1}{\epsilon} \left( \frac{a}{b} \right) \right\} > 1 - \sqrt{\epsilon} \quad (2)$$

*is greater than  $1 - \sqrt{\epsilon}$ .*

*Proof.* Let  $B$  be the complement of the set in (2). Then

$$\mathbb{E}\mu(B) = \int_Z \Pr(z \in B) dz = \int_Z \Pr(|\psi^{-1}(\psi(\phi(z))) \cap (\Phi(z) \setminus \{\phi(z)\})| > a/\epsilon b) dz.$$

Since the restriction of  $\psi$  to  $\Phi(z) \setminus \{\phi(z)\}$  is independent of  $\psi(\phi(z))$ ,

$$\mathbb{E}|\psi^{-1}(\psi(\phi(z))) \cap (\Phi(z) \setminus \{\phi(z)\})| = \frac{a}{b},$$

and hence the integrand is smaller than  $\epsilon$ . This means that  $\mathbb{E}\mu(B) < \epsilon$ , which implies the claim.  $\square$

### 3 Painting and Re-painting

From this point on, fix an ergodic system  $(Z, \mathcal{F}_Z, T)$ . All partitions are of  $Z$  and all sigma-algebras are contained in  $\mathcal{F}_Z$ .

**Definition 3.1.** A partition  $P$  is called  $\ell$ -admissible, if there is no  $z \in Z$  such that  $P(z) = P(Tz) = \dots = P(T^{\ell-1}z) = \bar{0}$ .

**Definition 3.2.** An  $\ell$ -admissible sequence of length  $n$  on  $a$  symbols is a sequence in  $\{\bar{0}, \dots, a-1\}^n$  whose first element is  $\bar{1}$ , and that does not contain a segment of consecutive  $\bar{0}$ 's of length  $\ell$ . The space of  $\ell$ -admissible sequences of length  $n$  on  $a$  symbols is denoted by  $\mathcal{A}(n, \ell, a)$ .

The following lemma is evident:

**Lemma 3.3.** For every  $\epsilon$ , if  $\ell$  is big enough, then  $\lim_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{A}(n, \ell, a)| > \log a - \epsilon$ .

This lemma will ensure that we can use the symbols  $\bar{0}$  and  $\bar{1}$  to mark the base of the tower and still have enough names for encoding.

**Definition 3.4.** Let  $(Z, \mathcal{F}_Z, \mu, T)$  be an ergodic system, and fix a generating partition  $\tilde{P}$ . Let  $\mathcal{T} = (A, M)$  be a tower, let  $\ell < |\mathcal{T}|$  be an integer, and let  $a \geq 2$  be an integer.

1. A  $(\mathcal{T}, \ell, a)$ -painting data is a function

$$\phi : |\tilde{P}|^{|\mathcal{T}|} \rightarrow \mathcal{A}(|\mathcal{T}| - \ell, \ell, a).$$

2. A painting data  $\phi$  induces a partition  $Q$  given by

$$z \mapsto \begin{cases} \phi(\nu_{T, \tilde{P}}(T^{-i}z))_i & T^{-i}z \in A \text{ and } 0 \leq i \leq |\mathcal{T}| - \ell - 1 \\ \bar{0} & \text{else} \end{cases}$$

**Lemma 3.5.** 1. For every tower  $\mathcal{T}$ , and every  $(\mathcal{T}, \ell, a)$ -painting data  $\phi$  that generates a partition  $Q$ , the base of  $\mathcal{T}$  is measurable with respect to  $\bigvee_{-\infty}^{\infty} T^i Q$ .

2. Suppose that  $h(\mathcal{T}) \geq \log a$  and let  $\epsilon > 0$ . There is  $\epsilon' > 0$  such that if  $\ell$  is big enough,  $\mathcal{T}$  is a tower that covers more than  $1 - \epsilon'$  of the space, and  $|\mathcal{T}|$  is big enough, then the probability that a random  $(\mathcal{T}, \ell, a)$ -painting data  $\psi$  that induces a partition  $Q$  satisfies that  $h(\mathcal{T}, Q) > \log a - \epsilon$ , is greater than  $1 - \epsilon$ .

*Proof.* The first claim follows, from the fact that  $z \in A$  if and only if  $P(z) = \bar{1}$  and  $P(T^{-1}z) = \dots = P(T^{-k}z) = \bar{0}$ .

As for the second claim, let  $\epsilon' < \epsilon$  be such that  $2\epsilon'|\tilde{P}| + H(\epsilon') < \epsilon$ , and assume that  $\mathcal{T}$  covers more than  $\epsilon'$  of the space. By Theorem 2.4, there is a collection  $\mathcal{N}$  of less than  $2^{(h(T)+\epsilon')|\mathcal{T}|}$   $(\mathcal{T}, P)$ -names such that the set of  $z \in A$  for which  $\nu_{\mathcal{T}, P}(z) \in \mathcal{N}$  has measure greater than  $(1 - \epsilon')\mu(A)$ . Applying Lemma 2.12 to the functions  $\Phi(z) = \mathcal{N}$  and  $\nu_{\mathcal{T}, P}(z)$ , we get that the probability that

$$\mu\{z \mid |\mathcal{N} \cap \psi^{-1}\psi(\nu_{\mathcal{T}, P}(z))| < 2^{(h(T) - \log a + \epsilon')|\mathcal{T}|}\} > 1 - \epsilon'$$

is greater than  $1 - \epsilon'$ . In this case, since  $z \mapsto \mathcal{N} \cap \psi^{-1}\psi(\nu_{\mathcal{T}, P}(z))$  is  $\bigvee T^i Q$ -measurable, Lemma 2.7 says that  $h(\mathcal{T} \mid \bigvee T^i Q) < h(T) - \log a + 2\epsilon'|\tilde{P}| + H(\epsilon') < h(T) - \log a + \epsilon$ , which implies the claim.  $\square$

**Definition 3.6.** Let  $(Z, \mathcal{F}_Z, \mu, T)$  be an ergodic system, and fix a generating partition  $\tilde{P}$ . Let  $P$  be a partition of  $Z$ , let  $\mathcal{T} = (A, M)$  be a tower, let  $\ell$  be an integer, and let  $\epsilon > 0$ .

1. A  $(\mathcal{T}, P, \ell, \epsilon)$ -repainting data is a function

$$\phi : |\tilde{P}|^{|\mathcal{T}|} \rightarrow \mathcal{A}(\epsilon|\mathcal{T}|, \ell, a).$$

2. A repainting data induces a partition of  $Z$  given by

$$z \mapsto \begin{cases} \bar{0} & z \notin \mathcal{T} \\ P(z) & T^{-i}z \in A, \epsilon|\mathcal{T}| \leq i \leq |\mathcal{T}| - 2\ell \\ \phi(\nu_{\mathcal{T}, P}(T^{-i}z))_i & T^{-i}z \in A, 0 \leq i \leq \epsilon|\mathcal{T}| \\ \bar{0} & T^{-i}z \in A, |\mathcal{T}| - 2\ell \leq i \leq |\mathcal{T}| \end{cases}$$

**Lemma 3.7.** Let  $P$  be a  $\ell$ -admissible partition, let  $\mathcal{T}$  be a tower that covers more than  $1 - \eta$  of the space, and let  $\phi$  be a  $(\mathcal{T}, P, \ell, \epsilon)$ -repainting data with associated partition  $Q$ . Then  $|P \triangle Q| < \epsilon + \eta$ .

**Lemma 3.8.** Let  $P$  be a  $\ell$ -admissible partition.

1. For every tower  $\mathcal{T} = (A, M)$ , if  $\phi$  is a  $(\mathcal{T}, P, \ell, \epsilon)$ -repainting data with associated partition  $Q$ , then  $A$  is measurable with respect to  $\bigvee_{-\infty}^{\infty} T^i Q$ .
2. If  $h(T, P) > \log a - \eta$ , then if  $\mathcal{T}$  is sufficiently invariant and covers more than  $1 - \eta$  of the space, then the probability that the partition  $Q$  associated with a random repainting data  $\psi$  satisfies  $h(T, Q) > \log a - \eta$  is greater than  $1 - \eta$ .

*Proof.* The first claim follows from the fact that  $z \in A$  if and only if  $P(z) = \bar{1}$  and  $Q(T^{-1}z) = Q(T^{-2}z) = \dots = Q(T^{-2\ell}z) = \bar{0}$ .

As for the second claim, define towers  $\mathcal{T}^{old} = (T^{\epsilon|\mathcal{T}|}A, (1 - \epsilon)|\mathcal{T}| - 2\ell)$  and  $\mathcal{T}^{new} = (A, \epsilon|\mathcal{T}|)$ . By Theorem 2.4, there is a collection  $\mathcal{N}$  of less than  $2^{(h(T) + \eta)|\mathcal{T}^{new}|}$   $(\mathcal{T}^{new}, P)$ -names such that the set of  $z \in A$  such that  $\nu_{\mathcal{T}^{new}, P}(z) \in \mathcal{N}$  has measure greater than  $(1 - \eta)\mu(z)$ .

Similarly to the proof of Lemma 3.5, there is a function  $\Phi : A \rightarrow 2^{\mathcal{N}_P}$  such that

1.  $\Phi(z)$  depends only on the  $(\mathcal{T}^{old}, P)$ -name of  $T^{\epsilon|T|}z$ .
2.  $|\Phi(z)| < 2^{(h(T,P) - \log a + \eta)|\mathcal{T}^{old}|}$ .
3. The set of  $z \in A$  such that  $\nu_{\mathcal{T}^{old}, P}(T^{\epsilon|T|}z) \in \Phi(z)$  has measure greater than  $(1 - \epsilon)\mu(A)$ .

By taking the product of  $\Phi$  and  $\mathcal{N}$ , there is a function  $\Psi : A \rightarrow 2^{\mathcal{N}_P}$  such that

1.  $\Psi(z)$  depends only on the  $(\mathcal{T}^{old}, Q)$ -name of  $T^{\epsilon|T|}z$ .
2.  $|\Psi(z)| < 2^{(h(T,P) - \log a + \eta)|\mathcal{T}^{old}| + h(T,P)|\mathcal{T}^{new}|}$ .
3. The set of  $z \in A$  such that  $\nu_{T,P}(z) \in \Psi(z)$  has measure greater than  $(1 - \epsilon)\mu(A)$ .

By Lemma 2.12, the probability that

$$\mu\{z \in A \mid |\Psi(z) \cap \phi^{-1}(\phi(\nu_{T,P}(z)))| < 2^{(h(T,P) - \log a + \eta)|\mathcal{T}^{old}| + (h(T,P) - \log a)|\mathcal{T}^{new}|}\} > (1 - \eta)\mu(A)$$

is greater than  $1 - \eta$ . In this case, since  $z \mapsto \Psi(z) \cap \phi^{-1}(\phi(\nu_{T,P}(z)))$  is  $\bigvee_{-\infty}^{\infty} T^i Q$ -measurable, Lemma 2.7 says that  $h(T, P \mid \bigvee T^i Q) < h(T, P) - \log a + \eta$ , which proves the claim.  $\square$

## 4 Proof of Theorem 1.2

Let  $(Z, \mathcal{F}_Z = \mathcal{F}_X \vee \mathcal{F}_Y, \mu, T)$  be as in Theorem 1.2. By Krieger's theorem, there are generating partitions  $\tilde{P}_X$  and  $\tilde{P}_Y$  for  $\mathcal{F}_X$  and  $\mathcal{F}_Y$  respectively. Let  $\epsilon_0 > 0$  be such that  $h(T) < \log a + \log b - \epsilon_0$ .

The first part of the proof of Theorem 1.2 is to obtain a pair of partitions that are close to generate, in the following sense:

**Definition 4.1.** A pair of partitions  $(P_X, P_Y)$  is called  $(\epsilon, \delta)$ -good if

1. The partition  $P_X$  is  $\mathcal{F}_X$ -measurable and has  $a$  parts; the partition  $P_Y$  is  $\mathcal{F}_Y$ -measurable and has  $b$  parts.
2.  $\tilde{P}_X \overset{\epsilon}{\subset} \bigvee_{-\infty}^{\infty} T^i(P_X \vee P_Y)$ , and  $\tilde{P}_Y \overset{\delta}{\subset} \bigvee_{-\infty}^{\infty} T^i(P_X \vee P_Y)$ .
3.  $h(T, P_Y) > \log b - \epsilon_0$ .

A pair of partitions  $(P_X, P_Y)$  is called  $(\epsilon, \delta)$ -very good if it is  $(\epsilon, \delta)$ -good and, in addition,  $h(T, P_X) > \log a - \epsilon_0$ .

Note that if an  $(\epsilon, \delta)$ -good pair exists, then  $h(T, \mathcal{F}_Y) > \log b - \epsilon_0$ . Similarly, if there is an  $(\epsilon, \delta)$ -very good pair, then also  $h(T, \mathcal{F}_X) > \log a - \epsilon_0$ .

**Proposition 4.2.** Assume that  $h(T, \mathcal{F}_Y) \geq \log b$  and  $h(T, \mathcal{F}_X) \geq \log a$ . Then for every  $\epsilon > 0$ , there is a pair of partitions  $(P_X, P_Y)$  which is  $(\epsilon, \epsilon)$ -very good.

*Proof.* Choose  $\eta > 0$  such that

$$h(T \mid \mathcal{F}_X) + 2\eta < \log b$$

and

$$2\eta(\log |\tilde{P}_Y| + 1) + H(\eta) < \epsilon_0$$



and

$$h(T) < \log a + \log b - H(\eta) - 2\eta(\log |\tilde{P}_Y| + 4).$$

Choose  $\ell$  such that

$$\log b - \eta < \lim_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{A}(n - \ell, \ell, b)|$$

and

$$\log a - \eta < \lim_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{A}(n - \ell, \ell, a)|.$$

1. *Choosing the Tower  $\mathcal{S}$* : By Theorem 2.4 and Corollary 2.6, if  $\mathcal{S}$  is an invariant enough tower with base  $A$  that covers more than half of the space, then there are

- A collection  $\mathcal{N} \subset \mathcal{N}_{\tilde{P}_Y}$  of size  $2^{|\mathcal{S}|(h(T, \mathcal{F}_Y) + \eta)}$ .
- A function  $\Phi_0$  from  $\mathcal{N}_{\tilde{P}_X}$  to subsets of  $\mathcal{N}_{\tilde{P}_Y}$  of size  $2^{|\mathcal{S}|(h(T, \mathcal{F}_X) + \eta)}$ .

such that the points  $z \in A$  for which

$$\nu_{\mathcal{S}, \tilde{P}_Y}(z) \in \mathcal{N} \quad (3)$$

and

$$\nu_{\mathcal{S}, \tilde{P}_Y}(z) \in \Phi_0(\nu_{\mathcal{S}, \tilde{P}_X}(z)) \quad (4)$$

has measure greater than  $(1 - \eta)\mu(A)$ . Choose such a tower  $\mathcal{S}$  that is measurable with respect to  $\mathcal{F}_Y$ , covers  $1 - \eta$  of the space, and such that  $|\mathcal{S}|$  satisfies

$$1 + \frac{1}{\eta} \left( \frac{2^{|\mathcal{S}|(h(T, \mathcal{F}_Y) + \eta)}}{|\mathcal{A}(|\mathcal{S}| - \ell, \ell, b)|} \right) < 2^{|\mathcal{S}|(h(T, \mathcal{F}_Y) - \log b + 2\eta)} \quad (5)$$

and

$$\frac{1}{\eta} \left( \frac{2^{|\mathcal{S}|(h(T, \mathcal{F}_X) + \eta)}}{|\mathcal{A}(|\mathcal{S}| - \ell, \ell, b)|} \right) < 1 \quad (6)$$

2. *Choosing the Partition  $P_Y$* : Choose a random function  $f : |\tilde{P}_Y|^{|\mathcal{S}|} \rightarrow \mathcal{A}(|\mathcal{S}| - \ell, \ell, b)$ , and let  $P_Y$  be the partition obtained by painting  $f$  on  $\mathcal{S}$ . It is clear that  $P_Y$  is  $\mathcal{F}_Y$ -measurable. Define a function  $\Phi_1$  from  $|P_Y|^{|\mathcal{S}|}$  to subsets of  $|\tilde{P}_Y|^{|\mathcal{S}|}$  by

$$\Phi_1(n) = \begin{cases} f^{-1}(n) \cap \mathcal{N} & |f^{-1}(n) \cap \mathcal{N}| < 2^{|\mathcal{S}|(h(T, \mathcal{F}_Y) - \log b + 2\eta)} \\ \emptyset & \text{else} \end{cases},$$

and a function  $\Psi$  from  $|\tilde{P}_X \times P_Y|^{|\mathcal{S}|}$  to  $|\tilde{P}_X \times \tilde{P}_Y|^{|\mathcal{S}|}$  by

$$\Psi(n, m) = \begin{cases} f^{-1}(m) \cap \Phi_0(n) & |f^{-1}(m) \cap \Phi_0(n)| = 1 \\ \text{undefined} & \text{else} \end{cases}.$$

By Lemma 2.12 and inequalities (5) and (6), most paintings satisfy that for all but  $2\eta$  of the points  $z$  in  $A$ ,

$$\nu_{\mathcal{S}, \tilde{P}_Y}(z) \in \Phi_1(\nu_{\mathcal{S}, P_Y}(z)), \quad (7)$$

and

$$\nu_{\mathcal{S}, \tilde{P}_X \times \tilde{P}_Y}(z) = \Psi(\nu_{\mathcal{S}, \tilde{P}_X \times P_Y}(z)). \quad (8)$$

By Lemma 2.7, we get that

$$h(T, \tilde{P}_Y | \bigvee T^i P_Y) < h(T, \mathcal{F}_Y) - \log b + 2\eta + H(\eta) + 2\eta \log |\tilde{P}_Y|. \quad (9)$$

In particular,  $h(T, P_Y) > \log b - 2\eta - H(\eta) - 2\eta \log |\tilde{P}_Y| > \log b - \epsilon_0$ .

3. *Choosing the Tower  $\mathcal{L}$* : By Corollary 2.6 and Equation (9), if  $\mathcal{L}$  is an invariant-enough tower with base  $B$ , then there is a function  $\Phi_2$  from  $|P_Y|^{|L|}$  to subsets of  $|\tilde{P}_Y|^{|L|}$  of size  $2^{|L|(h(T, \mathcal{F}_Y) - \log b + 2\eta(\log |\tilde{P}_Y| + 2) + H(\eta))}$  such that for all but  $\eta$ -portion of the points of  $B$ ,

$$\nu_{\mathcal{L}, \tilde{P}_Y}(z) \in \Phi_2(\nu_{\mathcal{L}, P_Y}(z)). \quad (10)$$

Also, by making  $\mathcal{L}$  more invariant and using Corollary 2.6, we can assume that there is a function  $\Phi_3$  from  $|\tilde{P}_Y|^{|L|}$  to subsets of  $|\tilde{P}_X|^{|L|}$  of size  $2^{|L|(h(T, \mathcal{F}_Y) + \eta)}$  such that for all but  $\eta$ -portion of the points  $z$  in  $B$ ,

$$\nu_{\mathcal{L}, \tilde{P}_X}(z) \in \Phi_3(\nu_{\mathcal{L}, \tilde{P}_Y}(z)). \quad (11)$$

By composing  $\Phi_2$  and  $\Phi_3$  we get a function  $\Phi_4$  from  $|P_Y|^{|L|}$  to subsets of  $|\tilde{P}_X|^{|L|}$  of sizes  $2^{|L|(h(T) - \log b + 2\eta(\log |\tilde{P}_Y| + 3) + H(\eta))}$  such that for all but  $2\eta$ -portion of the points  $z$  in  $B$ ,

$$\nu_{\mathcal{L}, \tilde{P}_X}(z) \in \Phi_4(\nu_{\mathcal{L}, P_Y}(z)). \quad (12)$$

Finally, we require that  $\mathcal{L}$  is  $\mathcal{F}_X$ -measurable and  $|\mathcal{L}|$  satisfies

$$\log |\mathcal{A}(|\mathcal{L}| - \ell, \ell, a)| > |\mathcal{L}|(\log a - \eta) \quad (13)$$

4. *Choosing the Partition  $P_X$* : Choose a random function  $g : |\tilde{P}_X|^{|L|} \rightarrow \mathcal{A}(|\mathcal{L}| - \ell, \ell, a)$ , and let  $P_X$  be the partition obtained by painting  $g$  on the tower  $\mathcal{L}$ . It is clear that  $P_X$  is  $\mathcal{F}_X$ -measurable. Define a function  $\Phi_5$  from  $|P_X \times P_Y|^{|L|}$  to  $|\tilde{P}_X \times P_Y|^{|L|}$  by

$$\Phi_5(n, m) = \begin{cases} (g^{-1}(n) \cap \Phi_4(m), m) & |g^{-1}(n) \cap \Phi_4(m)| = 1 \\ \text{undefined} & \text{else} \end{cases}$$

By Lemma 2.12, there is a function  $g$  as above such that for all but  $3\eta$  of the points in  $B$ ,

$$\nu_{\mathcal{L}, \tilde{P}_X \times P_Y}(z) = \Phi_5(\nu_{\mathcal{L}, P_X \times P_Y}(z)). \quad (14)$$

5. *Conclusion of the proof*: For all but  $2\eta$  of the points  $z \in Z$ ,  $z$  is in the image of both  $\mathcal{L}$  and  $\mathcal{S}$ . Assuming this is the case, by looking at the  $P_X$ -name of  $z$  we can find the smallest non-negative number  $i$  such that  $T^{-i}z \in B$ . Denote  $w = T^{-i}z$ . Similarly, the  $P_Y$ -name of  $z$  determines the minimal non-negative integer  $j$  such that  $T^{-j}z \in A$ . Denote  $u = T^{-j}z$ . The  $P_X \times P_Y$ -name of  $z$  determines  $\nu_{\mathcal{L}, P_X \times P_Y}(u)$ , and hence, for  $1 - 3\eta$  of the points, determines  $\nu_{\tilde{P}_X \times P_Y}(u) = \Phi_5(\nu_{\mathcal{L}, P_X \times P_Y}(u))$ . Let now  $j$  be the smallest non-negative number such that  $T^{-j}z \in A$ . The tuple  $\nu_{\tilde{P}_X \times P_Y}(u)$  determines  $\nu_{\mathcal{S}, \tilde{P}_X \times P_Y}(T^{-j}z)$  and so for  $1 - \eta$  of the points determines  $\nu_{\mathcal{S}, \tilde{P}_X \times \tilde{P}_Y}(T^{-j}z) = \Psi(\nu_{\mathcal{S}, \tilde{P}_X \times P_Y}(T^{-j}z))$ . This, clearly, determines  $\tilde{P}_X(z)$  and  $\tilde{P}_Y(z)$ .

□

The second part of the proof of Theorem 1.2 is to improve the pair  $(P_X, P_Y)$  and make it "more generating".

**Definition 4.3.** Given  $\epsilon, \delta > 0$ , let

$$f(\epsilon, \delta) = 2 \frac{H(\epsilon) + H(\delta) + \epsilon \log |\tilde{P}_X| + \delta \log |\tilde{P}_Y|}{\log a + \log b - \epsilon_0 - h(T)}.$$

Note that as  $\epsilon, \delta$  tend to 0, the function  $f(\epsilon, \delta)$  tends to 0.

**Proposition 4.4.** If  $\epsilon, \delta > 0$  are small enough then the following hold:

1. If  $(Q_X, Q_Y)$  is  $(\epsilon, \delta)$ -good, then for every  $\eta > 0$  there is a partition  $P_X$  such that  $(P_X, Q_Y)$  is  $(\eta, 2\delta)$ -good, and  $|P_X \triangle Q_X| < f(\epsilon, \delta)$ .
2. If  $(Q_X, Q_Y)$  is  $(\epsilon, \delta)$ -very good, then for every  $\eta > 0$  there is a partition  $P_X$  such that  $(P_X, Q_Y)$  is  $(\eta, 2\delta)$ -very good, and  $|P_X \triangle Q_X| < f(\epsilon, \delta)$ .

*Proof.* We can assume that  $\epsilon, \delta$ , and  $f(\epsilon, \delta)$  are less than  $1/2$ . There is a number  $N$  such that  $\tilde{P}_Y \subset \bigvee_{-N}^N T^i(Q_X \vee Q_Y)$ , and  $Q_X \subset \bigvee_{-N}^N T^i \tilde{P}_X$ . We can assume without loss of generality that  $4N\eta < \delta$  and that

$$f(\epsilon, \delta) > \frac{H(\epsilon) + H(\delta) + \epsilon \log |\tilde{P}_X| + \delta \log |\tilde{P}_Y| + \eta}{\log a + \log b - \epsilon_0 - h(T) - 2\eta}.$$

Choose  $\ell$  such that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{A}(n, \ell, a)| > \log a - \eta$$

By changing  $Q_X$  slightly (and enlarging  $\ell$  if needed), we can also assume that  $Q_X$  is  $\ell$ -admissible.

In order to construct the partition  $P_X$ , we choose a tower  $\mathcal{T} = (A, M)$  which is  $\mathcal{F}_X$ -measurable and covers more than  $1 - \eta$  of the space, and repaint the first  $f(\epsilon, \delta)|\mathcal{T}|$  levels of it using a random function  $\psi : |\tilde{P}_X|^{|\mathcal{T}|} \rightarrow \mathcal{A}(\ell, f(\epsilon, \delta)|\mathcal{T}|, a)$ . We will show that if  $\mathcal{T}$  is taken as sufficiently invariant, then with high probability (on  $\psi$ ), the obtained partition—which we denote by  $P_X$ —is good. By Lemma 3.7,  $|P_X \triangle Q_X| < f(\epsilon, \delta) + \eta$ .

Given  $\mathcal{T}$ , define the towers  $\mathcal{T}^{old} = (T^{f(\epsilon, \delta)|\mathcal{T}|} A, (1 - f(\epsilon, \delta))|\mathcal{T}| - 2\ell)$  and  $\mathcal{T}^{new} = (A, f(\epsilon, \delta)|\mathcal{T}|)$ . As  $|\mathcal{T}| \rightarrow \infty$ , the towers  $\mathcal{T}^{old}$  and  $\mathcal{T}^{new}$  become more and more invariant. Note that both cover more than  $f(\epsilon, \delta)/2$  of the space if  $|\mathcal{T}|$  is invariant enough, and that  $\nu_{\mathcal{T}^{old}, P_X}(z) = \nu_{\mathcal{T}^{old}, Q_X}(z)$  for every  $z \in T^{f(\epsilon, \delta)|\mathcal{T}|} A$ .

Applying Proposition 2.11 to the pairs  $(\tilde{P}_X, Q_X \vee Q_Y)$  and  $(\tilde{P}_Y, Q_X \vee Q_Y)$ , if  $\mathcal{T}$  is invariant enough, then there is a function  $\Phi_1 : \mathcal{N}_{Q_X \vee Q_Y} \rightarrow 2^{\mathcal{N}_{\tilde{P}_X \vee \tilde{P}_Y}}$  such that for any  $z \in T^{f(\epsilon, \delta)|\mathcal{T}|} A$ , the set  $\Phi_1(\nu_{\mathcal{T}^{old}, Q_X \vee Q_Y}(z))$  has at most

$$2^{(H(\epsilon) + H(\delta) + \epsilon|\tilde{P}_X| + \delta|\tilde{P}_Y| + \eta)|\mathcal{T}|}$$

elements, and the set of  $z \in T^{f(\epsilon, \delta)|\mathcal{T}|} A$  for which  $\nu_{\mathcal{T}^{old}, \tilde{P}_X \vee \tilde{P}_Y}(z) \in \Phi_1(\nu_{\mathcal{T}^{old}, Q_X \vee Q_Y}(z))$  has measure larger than  $(1 - \eta)\mu(A)$ .

By Corollary 2.6 applied to the pair  $(Q_Y, \tilde{P}_Y)$ , if  $\mathcal{T}$  is invariant enough, then there is a function  $\Phi_2 : \mathcal{N}_{Q_Y} \rightarrow 2^{\mathcal{N}_{\tilde{P}_Y}}$  such that for any  $z \in A$ , the set  $\Phi_2(\nu_{\mathcal{T}^{new}, Q_Y}(z))$  has at most

$$2^{(h(\mathcal{T}, \mathcal{F}_Y) - \log b + \epsilon_0 + \eta)|\mathcal{T}^{new}|}$$

elements, and the set of  $z \in A$  such that  $\nu_{\mathcal{T}^{new}, \tilde{P}_Y}(z) \in \Phi_2(\nu_{\mathcal{T}^{new}, Q_Y}(z))$  has measure greater than  $(1 - \eta)\mu(A)$ .

Applying Corollary 2.6 to the pair  $(\tilde{P}_X, \tilde{P}_Y)$ , if  $\mathcal{T}$  is invariant enough, then there is a function  $\Phi_3 : \mathcal{N}_{\tilde{P}_Y} \rightarrow 2^{\mathcal{N}_{\tilde{P}_X}}$  such that for every  $z \in A$ , the set  $\Phi_3(\nu_{\mathcal{T}^{new}, \tilde{P}_Y}(z))$  has at most

$$2^{(h(\mathcal{T}|\mathcal{F}_Y) + \eta)|\mathcal{T}^{new}|}$$

elements, and the set of  $z \in A$  such that  $\nu_{\mathcal{T}^{new}, \tilde{P}_X}(z) \in \Phi_3(\nu_{\mathcal{T}^{new}, \tilde{P}_Y}(z))$  has measure greater than  $(1 - \eta)\mu(A)$ .

Let  $\Phi_4$  be the composition of  $\Phi_3$  and  $\Phi_2$ . Combining  $\Phi_1$  and  $\Phi_4$ , we get a function  $\Phi_5 : \mathcal{N}_{Q_X} \times \mathcal{N}_{Q_Y} \rightarrow 2^{\mathcal{N}_{\tilde{P}_X}}$  such that for every  $z \in A$ , the set  $\Phi_5(\nu_{\mathcal{T}^{old}, Q_X}(T^{f(\epsilon, \delta)}|Tz), \nu_{\mathcal{T}, Q_Y}(z))$  has size at most

$$2^{(H(\epsilon) + H(\delta) + \epsilon|\tilde{P}_X| + \delta|\tilde{P}_Y| + f(\epsilon, \delta)(h(\mathcal{T}) - \log b + \epsilon_0) + 3\eta)|\mathcal{T}|},$$

and the set of  $z \in A$  such that

$$\nu_{\mathcal{T}, \tilde{P}_X}(z) \in \Phi_5(\nu_{\mathcal{T}^{old}, Q_X}(T^{f(\epsilon, \delta)}|Tz), \nu_{\mathcal{T}, Q_Y}(z))$$

has measure greater than  $(1 - 2\eta)\mu(A)$ .

Note that if  $|\mathcal{T}|$  is large enough, then

$$\frac{2^{(H(\epsilon) + H(\delta) + \epsilon|\tilde{P}_X| + \delta|\tilde{P}_Y| + f(\epsilon, \delta)(h(\mathcal{T}) - \log b + \epsilon_0) + 2\eta)|\mathcal{T}|}}{|\mathcal{A}(f(\epsilon, \delta)|\mathcal{T}|, \ell, a)|} <$$

$$= 2^{(H(\epsilon) + H(\delta) + \epsilon|\tilde{P}_X| + \delta|\tilde{P}_Y| + f(\epsilon, \delta)(h(\mathcal{T}) - \log b - \log a + \epsilon_0) + 3\eta)|\mathcal{T}|} \rightarrow 0$$

as  $|\mathcal{T}| \rightarrow \infty$ . Applying Lemma 2.12 to the functions  $\Phi_5(\nu_{\mathcal{T}^{old}, Q_X}(T^{f(\epsilon, \delta)}|Tz), \nu_{\mathcal{T}, Q_Y}(z))$ ,  $\nu_{\mathcal{T}, \tilde{P}_X}(z)$ , and the random coloring  $\psi$ , we get that with high probability on  $\psi$ , the set of points  $z \in A$  for which the intersection

$$\Phi_6(z) = \Phi_5(\nu_{\mathcal{T}^{old}, Q_X}(T^{f(\epsilon, \delta)}|Tz), \nu_{\mathcal{T}, Q_Y}(z)) \cap \psi^{-1}\psi(\nu_{\mathcal{T}, \tilde{P}_X}(z))$$

is a singleton has measure greater than  $(1 - 3\eta)\mu(A)$ .

We claim that

$$\tilde{P}_X \subset \bigvee_{-\infty}^{4\eta} T^i(P_X \vee Q_Y).$$

Indeed, let  $\mathcal{G} = \bigvee_{-\infty}^{\infty} T^i(P_X \vee Q_Y)$ . By Lemma 3.8, the set  $A$  is  $\mathcal{G}$ -measurable, and so is the function  $\beta : Z \rightarrow \mathbb{Z}$  that sends  $z \in Z$  to the largest non-positive integer  $i$  such that  $T^i z \in A$ . The  $\mathcal{G}$ -measurable partition

$$z \mapsto \begin{cases} (\Phi_6(T^{\beta(z)} z))_{-\beta(z)} & \text{if } \beta(z) > -M \text{ and } \Phi_6(z) \text{ is a singleton} \\ \bar{0} & \text{else} \end{cases}$$

coincides with  $\tilde{P}_X$  for all but  $1 - 4\eta$  of the space.

As for  $\tilde{P}_Y$ , we have that

$$\tilde{P}_Y \subset \bigvee_{-N}^N T^i Q_X \vee \bigvee_{-N}^N T^i Q_Y \stackrel{2N\eta}{\subset} \bigvee_{-2N}^{2N} T^i \tilde{P}_X \vee \bigvee_{-N}^N T^i Q_Y \stackrel{4N\eta}{\subset} \mathcal{G}$$

and so

$$\tilde{P}_Y \stackrel{\delta+6N\eta}{\subset} \mathcal{G}.$$

Finally, if  $h(T, Q_X) > \log a$ , then by Lemma 3.8, if  $\mathcal{T}$  is sufficiently invariant, the entropy  $h(T, P_X)$  is greater than  $\log a$  with high probability on the random function  $\psi$ .  $\square$

*Proof of Theorem 1.2.* If  $h(T, \mathcal{F}_X) < \log a$  and  $h(T, \mathcal{F}_Y) < \log b$ , then by Krieger's theorem there are partitions  $\tilde{P}_X$  and  $\tilde{P}_Y$  that are  $\mathcal{F}_X$  and  $\mathcal{F}_Y$  measurable, have  $a$  and  $b$  parts, and generate  $\mathcal{F}_X$  and  $\mathcal{F}_Y$  respectively. We assume in the following that this does not hold.

Choose  $\xi_0$  such that Proposition 4.4 applies to the pair  $(2\xi_0, 2\xi_0)$ . Inductively choose a decreasing sequence  $\xi_n$  such that both the sums  $\sum_n f(\xi_n, \xi_n)$  and  $\sum_n f(\xi_{n+1}, 2\xi_n)$  converge, where  $f$  is defined in 4.3.

If  $h(T, \mathcal{F}_X) < \log a$  and  $h(T, \mathcal{F}_Y) \geq \log b$ , then by Krieger's theorem, there is a generating partition  $\tilde{P}_X$  to  $\mathcal{F}_X$ . We define a sequence of partitions  $P_Y^n$  as follows: by Proposition 4.2, there is a pair of partitions  $(P_X, P_Y^0)$  that is  $(\xi_0, \xi_0)$ -good. It follows that  $(\tilde{P}_X, P_Y^0)$  is also  $(0, \xi_0)$ -good. Assuming  $P_Y^n$  was defined, applying Proposition 4.4, there is a partition  $P_Y^{n+1}$  such that  $(\tilde{P}_X, P_Y^{n+1})$  is  $(0, \xi_{n+1})$ -good and  $|P_Y^n \triangle P_Y^{n+1}| < f(0, \xi_n) < f(\xi_n, \xi_n)$ . Therefore  $\sum_n |P_Y^n \triangle P_Y^{n+1}| < \infty$ , so there is a limit partition  $P_Y^\infty$ . It follows that  $(\tilde{P}_X, P_Y^\infty)$  satisfy the requirements of the theorem. The same proof holds if  $h(T, \mathcal{F}_Y) < \log b$  and  $h(T, \mathcal{F}_X) \geq \log a$ .

Assume finally that  $h(T, \mathcal{F}_X) \geq \log a$  and  $h(T, \mathcal{F}_Y) \geq \log b$ . We define a sequence of pairs of partitions  $(P_X^n, P_Y^n)$  as follows: by Proposition 4.2, there is a pair of partitions  $(P_X^0, P_Y^0)$  that is  $(\xi_0, \xi_0)$ -very good. Assuming  $(P_X^n, P_Y^n)$  have been defined and the pair is  $(\xi_n, \xi_n)$ -very good, applying Proposition 4.4 there is a partition  $P_X^{n+1}$  such that  $(P_X^{n+1}, P_Y^n)$  is  $(\xi_{n+1}/2, 2\xi_n)$ -very good and  $|P_X^n \triangle P_X^{n+1}| < f(\xi_n, \xi_n)$ . Applying Proposition 4.4 again, there is a partition  $P_Y^{n+1}$  such that the pair  $(P_X^{n+1}, P_Y^{n+1})$  is  $(\xi_{n+1}, \xi_{n+1})$ -very good and  $|P_Y^n \triangle P_Y^{n+1}| < f(\xi_{n+1}, 2\xi_n)$ . Since by definition of the  $\xi_n$ 's, the sums  $\sum |P_X^n \triangle P_X^{n+1}|$  and  $\sum |P_Y^n \triangle P_Y^{n+1}|$  converge, there are limit partitions  $P_X^\infty$  and  $P_Y^\infty$ . The pair  $(P_X^\infty, P_Y^\infty)$  satisfies the requirements of the theorem.  $\square$

## References

- [1] Avni, N. Entropy Theory for Cross Sections. preprint. available at <http://arxiv.org/abs/math/0609273>.
- [2] Gray, R. M.; Ornstein, D. S.; Dobrushin, R. L. Block synchronization, sliding-block coding, invulnerable sources and zero error codes for discrete noisy channels. Ann. Probab. 8 (1980), no. 4, 639–674.

- [3] Kieffer, J. Zero-error stationary coding over stationary channels. *Z. Wahrsch. Verw. Gebiete* 56 (1981), no. 1, 113–126.
- [4] Krieger, W. On generators in ergodic theory, *Internat. Congress Math. (Vancouver, B.C., 1974)*, vol. 2, *Canad. Math. Congress*, 1975, pp. 303–308.
- [5] Rudolph, D. *Fundamentals of Measurable Dynamics, Ergodic Theory on Lebesgue Spaces*, 1990, Oxford University Press.
- [6] Slepian, D.; Wolf, J.K. Noiseless coding of correlated information sources. *IEEE Trans. Information Theory* IT-19 (1973), 471–480.